# REPORT DOCUMENTATION PAGE

Form Approved
OMB NO. 0704-0188

Public Reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comment regarding this burden estimates or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188,) Washington, DC 20503.

| 1. AGENCY USE ONLY ( Leave Blank) | 2. REPORT DATE<br>12/21/2000 | 3. REPORT TYPE AND DATES COVERED<br>Final Progress Report  8/12/96-6/20/2000 |
|---|---|---|

| 4. TITLE AND SUBTITLE<br><br>Scalable and Robust Video Compression | 5. FUNDING NUMBERS<br><br>DAAH04-96-1-0380 |
|---|---|

**6. AUTHOR(S)**
John W. Woods

| 7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)<br><br>ECSE Department, Rensselaer Polytechnic Institute, Troy, NY 12180-3590 | 8. PERFORMING ORGANIZATION REPORT NUMBER |
|---|---|

| 9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)<br><br>U. S. Army Research Office<br>P.O. Box 12211<br>Research Triangle Park, NC 27709-2211 | 10. SPONSORING / MONITORING AGENCY REPORT NUMBER<br><br>ARO 35081.5-CI |
|---|---|

**11. SUPPLEMENTARY NOTES**
    The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.

| 12 a. DISTRIBUTION / AVAILABILITY STATEMENT<br><br>Approved for public release; distribution unlimited. | 12 b. DISTRIBUTION CODE |
|---|---|

**13. ABSTRACT (Maximum 200 words)**

The research tasks are: Task 1, *Scalable Coding*; Task 2, *Alternatives to variable length coding* (VLC); and Task3, *Packetizing for Error Concealment*. Initial work on Task 1 built on research from a prior DARPA contract, and resulted in a scalable video coder with the unique property of scalable transmission at six resolution/frame-rate combinations. Task 2 work combined fixed and variable length coding (VLC) into a joint coding approach with distinct robustness advantages, while giving up only a slight amount of performance versus straight VLC coding in the noise-free case.

Task 3 *Packetzing for Error Concealment* explored the new technique *dispersive packetization* and combined it with error concealment in a state-of-the-art subband/wavelet image coder. An *ns-2* network simulation revealed substantial robustness to packet loss error, with useable images recovered under packet loss rates at high as 30%, thus implementing the graceful degradation property.

QUALITY INSPECTED
20010117 128

| 14. SUBJECT TERMS | | | 15. NUMBER OF PAGES |
|---|---|---|---|
| | | | 16. PRICE CODE |

| 17. SECURITY CLASSIFICATION OR REPORT<br>**UNCLASSIFIED** | 18. SECURITY CLASSIFICATION ON THIS PAGE<br>**UNCLASSIFIED** | 19. SECURITY CLASSIFICATION OF ABSTRACT<br>**UNCLASSIFIED** | 20. LIMITATION OF ABSTRACT<br><br>**UL** |
|---|---|---|---|

NSN 7540-01-280-5500

Standard Form 298 (Rev.2-89)
Prescribed by ANSI Std. 239-18
298-102

# Scalable and Robust Video Compression

John W. Woods

Center for Image Processing Research and

ECSE Department

Rensselaer Polytechnic Institute

Troy, NY 12180-3590

# I.    Problems Studied

We studied digital image and video compression from the viewpoint of both scalability and robust transmission. Scalability means that the image (video) can be coded just once and then decoded at various resolutions and quality levels for transmission in an efficient manner. In the case of video, we add frame rate to the list of desirable scalability types. Robustness means that the compressed data can be transmitted in a hostile environment consisting of packet losses (wired case) and bit errors (wireless case).

We completed development of a scalable subband/wavelet coder offering three spatial resolutions and two frame rates. This transform based, non-hybrid coder, had both scalable and robustness qualities. We then turned to look at an innovative combination of fixed and variable-length coding to protect the output of a DCT-type image coder. In the last year of the grant, we investigated our method using dispersive packetization and error concealment to reduce the effect of packet losses. We have done our work using UC Berkely's *ns-2* network simulator.

# II.    Most Important Results

The most important result from Task 1 was completion of the transform-based subband/wavelet coder with six combinations of resolution and frame rate, all at efficient compression level of 0.35 bits per pixel.

The most important result from Task 2 was the demonstration that a combination of fixed-plus-variable length coding can give a very robust performance at high bit error rates.

The most important result from Task 3 is showing through network simulation that dispersive packetization combined with the type of error concealment permitted by subband/wavelet coders can lead to graceful degradation with increasing packet loss ratio, all the way up to 30%.

# III.    More Detailed Results

The research tasks were: Task 1, *Scalable Coding*; Task 2, *Alternatives to variable length coding*; and Task3, *Packetizing for Error Concealment*.

Initially we concentrated on Task 1, *Scalable Coding,* where existing doctoral student Gary Lilienfield completed his PhD thesis. Lilienfield succeeded in going beyond scalable coding to achieve scalable transmission for the test videos considered. By this we mean that not only is the compression algorithm scalable, but it also permitted sharing of transmission links in an efficient manner. His coder is able to combine temporal and spatial scalability to save bandwidth on simultaneous digital network

transmission to two or more receivers requesting different resolutions and/or frame rates. Currently the system has two frame rates and three spatial resolutions, giving a total of six combinations. Efficiency is comparable to state-of-the-art coders without scalability, with only a 0.5 dB penalty averaged over the six 'scales.'

Work during 1997, the first full year of the grant, concentrated on Task 2, *Alternatives to variable length coding*. This effort combined fixed and variable length coding and was largely the work of Dr. Aydin Alatan, a postdoctoral assistant supported by the grant in the second year. This work has shown that combinations of fixed and variable length codes can be very effective for combating noise for the block DCT image coding method (cf. JPEG) that we used and for the binary symmetric test channel employed. We compared our fixed-plus-variable source coder with standard variable length coders with synch words. We found that the new method outperformed the standard variable length method at bit-error rates (BER) of $10^{-1}$, $10^{-2}$, and $10^{-3}$ for our test images, sometimes by as much as 14 dB. This is even after optimizing the slice length to give the variable length codes the best chance to re-synchronize.

In this work we did not employ error correction coding (ECC), because to first order, ECC just lowers the BER of the channel, but at the expense of a lower information transmission rate. Thus our BER $10^{-3}$ results can correspond to a binary symmetric channel plus ECC, which has a BER of $10^{-3}$ *after error correction*. Of course, nonstationary effects, such as encountered in bursty channels, go beyond this simple assumption, and would need to be studied further. This would also not follow for protection against packet loss, which was not considered in this Task.

Dr. Alatan completed his work on this project by concentrating on the extension of combined fixed and variable length coding to then state-of-the-art subband/wavelet coders, i.e. SPIHT, layered zero coding (LZC), and wavelet trellis coded quantization (WTCQ). Unlike the conventional wavelet-based methods, the lowest subband was encoded by fixed-length (or fixed and variable-length) to achieve better immunity to channel synchronization errors. In such a scheme, the receiver is able to reconstruct a coarser version of the image even in the case of channel mismatches, which can easily distort error-protected variable-length data. Fixed-length coding was achieved using Lloyd-Max Quantization. Other subbands were coded in an embedded manner using a scheme based on our hybrid method. In this scheme, while significant coefficients are progressively encoded in "unsaturated" regions, the saturating values are variable-length coded at a final stage. In order to encode the locations of significant coefficients, zero trees (or set partitions) were utilized and the bit-stream carrying this location information is protected separately.

Our novel approach of combining fixed and variable length coding is seen to enjoy a robustness advantage against variable length coding alone, and a quality (or efficiency) advantage against fixed length coding a one. It seems that the PSNR advantage of the technique lies in a transitional BER region, where the VLC part still helps, dependent on its FEC coding. When the BER would jump higher, of course, there is the *robustness advantage* of getting *something recognizable* through the fixed length channel.

Most work in the third year 1999 was devoted to Task 3, *Packetizing for Error Concealment.*
Back in 1995, the proposal for this grant had described a then novel approach to packetizing video
for transmission on lossy (packet lossy) networks. However since that time similar ideas have
appeared in the literature. Our overall coder first decomposes an image or each frame of a video into
a critically sampled subband/wavelet pyramid. Then it quantizes and packetizes the resulting
coefficients (samples) in such a way that loosing one packet will only cause the loss of a single
subband/wavelet coefficient in any *local* neighborhood, both in space and in frequency. As mentioned
above, we call the process that does this assignment *dispersive packetization.* Our research, and
that of others by now, has shown that the concept is quite valid for normal sized images, their
compressibilities, and current network-friendly packet sizes (around 550 Bytes to avoid
fragmentation). *Dispersion in space* allows interpolation from nearest neighbor coefficients to cover
for a missing coefficient. *Dispersion in frequency* allows utilization of the nonlinear dependence
between nearest neighbor frequency coefficients at the same location. Channel coding can do a similar
dispersion of the bit stream, while we are implementing our dispersion *at the symbol level.* In channel
coding literature, such *interleavers* are used for burst noise and erasure channels. By contrast, our
dispersive packetizers are designed to separate local coefficients and put them in different packets,
with the separation done before any entropy or variable length coding.

Dispersive packetization can be viewed as a type of *Multiple Description* coding, whose stated
purpose is to deal with packet lossy channels by making, ideally, all packets equally important. In
conventional multiple description coding, this is done with two (or more) embedded quantizers, say
one outputting the even numbers and one the odd numbers. Then when the outputs of the even and
odd quantizers are sent as separate packets, each provides a coarse description of the input data. If
both packets are received, then the full accuracy can be achieved. Dispersive packetization achieves
a similar end of making all the packets approximately equally important, and provides a coarse
reproduction, showing interpolation error in place of quantization error, when a packet is lost. Some
combination of these ideas may be useful.

We have compared dispersive packetization (DP) with regular or *block packetization* (BP) for the
test sequence *football*, with both error concealment and no error concealment. We find that, at a
25% average packet loss rate, error concealment can raise the PSNR for DP by about 5 dB and for
BP by about 4.2 dB. However, when we actually look at the pictures, we see large blurred
'information gaps' in the BP based frames, where the interpolation has succeeded in reproducing only
the lowest frequencies. While in the DP based images the frames are of a much more uniform quality,
with the interpolation needing to be performed over a much reduced spatial area. Thus DP with error
concealment does not produce any severe 'information gaps' in the decoded frames.

Our network simulations have used the *ns-2* software. We transmit in an RTP/UDP method, thus
avoiding the slowdown and slow startup of TCP, generally in use on the public Internet for transmitting
images and data. In our simulations, and as others have found too, RTP performed much better that
TCP with bit rates in the vicinity of 40 Kbps, while TCP only achieved around 5 Kbps. We seek to
avoid fragmentation at internal network nodes by packetizing into roughly fixed size packets around

550 KB in size. We experimented with several network configurations, most with low or negligible bit error rate (BER). However, we did extend the simulation to small ring subnets with relatively high $10^{-2}$ BER more typical of tactical radio nets, with good results.

## Future Plans

We planed to investigate incorporation of motion compensation and also at least two levels of priority or protection into our dispersive packetization scheme. We expected to improve our embedded coder somewhat over the currently used Dartmouth coder, by including a new state-of-the-art coder developed at Rensselaer *Embedded Zero-Block Coder* (EZBC), which should be capable of simultaneous scalable frame rate and resolution performance, as required in a heterogeneous and time-varying network.

# IV. List of Manuscripts

## Conference Papers

1. "Scalable High Definition Video Coding," by Gary Lilienfield and John W. Woods, *Proc. SPIE Visual Communications and Image Processing*, VCIP-98, San Jose, CA, January 1998.

2. "Joint Utilization of Fixed and Variable-length Codes for Improving Synchronization Immunity for Image Transmission, " *Proc. IEEE International Conference on Image Processing,* ICIP-98, Chicago, IL, October 1998.

3. "Robust Signal Compression Using Fixed- and Variable-Length Coding," by A. Aydin Alatan and John W. Woods, *ARMY Workshop on Data Compression Processing Techniques for Missile Guidance Data Links*, Redstone Arsenal, AL, December 1998.

4. "Robust transmission of packet video through dispersive packetization and error concealment," by Ivan V. Bajic, John W. Woods, and Ali Mohammad Chaudry, *Proc. Packet Video Workshop* (PV2000), Cagliari, Italy, May 2000.

5. "Robust Signal Compression using Joint Fixed- and Variable-length Coding," by A. Aydin Alatan and John W. Woods, *Proc. International Symposium on Information Theory* (summaries only) , ISIT-2000), Sorrento, Italy, June 2000.

## Manuscripts

1. "Robust Signal Compression using Joint Fixed- and Variable-length Coding," by A. Aydin Alatan and John W. Woods, in revision.

2. "A Subband/Wavelet Resolution and Frame-Rate Scalable Video Coder," by John W. Woods and Gary Lilienfield, submitted to *IEEE Trans. CSVT*, in revision.

Report

1. *Zero-tree based Fixed-length Image Coding for Noisy Channels* by A. Aydin Alatan and John W. Woods, Center for Image Processing Research (CIPR), Rensselaer Polytechnic Institute, Troy, NY, August 1998.

# V.   Scientific Personnel

Principal Investigator: John W. Woods
Post-Doctoral Associate: A. Aydin Alatan
Research Assistants: Gary Lilienfield (PhD 1997), Suleyman Kozat, Ivan Bajic (MS June 2000), and Ali M. Chaudry (MS December 2000)

# VI.   Report of Inventions

We regard our method *Dispersive Packetization* as an invention, but have no present plans to proceed with patenting.